



**University of
Zurich**^{UZH}

**Zurich Open Repository and
Archive**

University of Zurich
University Library
Strickhofstrasse 39
CH-8057 Zurich
www.zora.uzh.ch

Year: 2018

Virtual and augmented realities – a conversation with William Uricchio

Bornkamm, Henriette ; Köhler, Kristina ; Petraitis, Marian

Abstract: Warum Virtual und Augmented Reality aktuell in aller Munde sind, was für Herausforderungen und Möglichkeiten die neuen Technologien mit sich bringen und inwiefern es sich dabei wirklich um die Zukunft des Kinos handelt, haben Henriette Bornkamm, Kristina Köhler und Marian Petraitis mit dem US-Medienwissenschaftler William Uricchio für CINEMA besprochen.

Posted at the Zurich Open Repository and Archive, University of Zurich
ZORA URL: <https://doi.org/10.5167/uzh-162725>
Journal Article
Published Version

Originally published at:

Bornkamm, Henriette; Köhler, Kristina; Petraitis, Marian (2018). Virtual and augmented realities – a conversation with William Uricchio. *Cinema*, 63:38-49.

H. BORNKAMM, K. KÖHLER, M. PETRAITIS VIRTUAL AND AUGMENTED REALITIES — A CONVERSATION WITH WILLIAM URICCHIO

The future of cinema has been subject of controversial debates in different phases of film history. In 1946, film theorist André Bazin articulated his idea of 'Total Cinema', where reality would be replaced by «a perfect illusion of the outside world in sound, colour, and relief».¹ This idea is currently experiencing a new hype, especially renewed by practices of Virtual and Augmented Reality (shortened VR and AR). And yet, the promises and claims connected to those terms are not new at all. In the 1990s, the singer Jamiroquai already bewailed the «virtual insanity that we're livin' in», urging us to leave the virtual world as fast as possible; his rather moralistic message was highly double-coded back then: In the music video to «Virtual Insanity», we see the singer beaming himself and his furniture through a conspicuously futuristic interior. Since then, VR-technology has merged into mainstream culture: from TV stations to super markets, companies are using do-it-yourself VR-headsets to advertise their shows and products, bringing VR into private households. Augmented reality games such as «Pokémon Go» are introducing the new technology via the entertainment industry to larger audiences.

But what exactly do we mean when we talk of Virtual and Augmented Reality? What are the implicit und explicit promises of these seemingly new phenomena? And how exactly are the new technologies used in practice?

William Uricchio, a US media scholar and Professor of Comparative Media Studies at the Massachusetts Institute of Technology and of Comparative Media History at the Universiteit Utrecht, has been dealing with these questions in his research for several years already. At MIT, he leads the Open Documentary Lab that brings together storytellers, technologists, and scholars to explore how these new technologies are transforming our media practices –

and especially documentary formats. In June, Uricchio was invited as keynote speaker to a workshop of the Department of Film Studies at the University of Zurich. Henriette Bornkamm, Kristina Köhler and Marian Petraitis met with him for CINEMA Jahrbuch to talk about Virtual and Augmented Reality, and the possible future(s) of cinema.



William Uricchio

CINEMA: Virtual Reality and Augmented Reality are current buzzwords in the world of film and beyond. This hype around the visionary power of VR and AR seems to raise a number of questions, such as how to think about the future of film, media and representation, about the viewer and the author/creator, about experience and agency, about our idea of the visual image and how it is created. Before we get into more details with these terms, we would like to start with a polemical question that a lot of people might have in mind when they first experience VR/AR: Is this still film? And if so: Is this the future of film – or one of its many possible futures?

WILLIAM URICCHIO: Two good questions! Film is a term that can be defined narrowly («film and media studies», as if film were somehow not a medium) or more broadly, including motion pictures streamed on computer screens and made-for-television movies. Whether or not the rubric film includes VR depends in part on just how broad one's definition is, or what one's definition

privileges (moving images? sites of exhibition? production process?). In an era when film organizations are often part of larger multi-media companies, when filmic assets slide effortlessly across platforms, and when cross- and trans-media productions are increasingly the norm, these definitional constructions are dynamic. Consider the rise of «VR cinemas», or the growing presence of VR at film festivals like Sundance and Tribeca, or Netflix's VR «movies» or Chris Milk's «landmark VR film», *Evolution of Verse* (US 2015), and it's clear at least that some prominent makers, distributors and exhibitors are using the terms film and VR in one breath.



Photo-realistic, CGI-rendered Virtual Reality in *Evolution of Verse* (Chris Milk, USA 2015)

Is this the future of film? I doubt it. But it certainly could be a future, among many others ... a delivery platform ... or an element in a larger ensemble. In answering this question, I'm aware of standing on the cusp of a new technological era, I'm aware of André Bazin's important provocations in «The Myth of Total Cinema»² with its easy-to-extend implications for aspects of VR. And I'm aware of my personal deep-seated prejudice to the effect that films should be capable of being seen with a collective audience (just like the prejudice that television should be capable of liveness). In fact, precious little TV is live, and I see too many films in empty cinemas ... as I said, «prejudice!» But VR currently tends to be isolating, much like novel reading, and so while granting that it might be a future of film, it's a future that largely rubs against the grain of 20th century configurations of film.

CINEMA: *Maybe at this point, it could be helpful to specify what we mean by Virtual and Augmented Reality (AR). What falls under the category Virtual Reality and how can it be distinguished from Augmented Reality?*

URICCHIO: VR is currently a term that is used with abandon! It can refer to technologies as diverse as CAVEs (cave automatic virtual environments)³, 360° video, real time VR (laser-scanned data points and activating algorithms), CGI (computer generated imagery), and more. I even saw a recent Kickstarter campaign for a VR viewing system that uses «Pepper's Ghost», a 19th century angled-glass system for creating stage illusions. These various systems can create a sense of immersion simply by offering a world to look at from a fixed position, or more complexly, by creating interactive experiences with that world in which the user can walk around. And if the past is prologue, humans have a remarkable capacity to move from shock to nonchalance with repeated exposures to the same stimulus. So, whatever «works» convincingly as virtually real today will probably look slightly contrived in a few years.

Virtual reality today generally refers to a computer-generated emulation of experience, often in the form of a world that the user can interact with and explore. It can use visual, haptic, acoustic, and even olfactory cues to emulate a world that is realistic ... or not. There is plenty of room for ambiguity, particularly at a moment when, at one end of the spectrum, the simplicity of 360° video has led to widespread use (the *New York Times* has a daily 360° feature, and YouTube has a 360° channel); and at the other end, systems that enable true interaction with their worlds (LIDAR⁴, Kinect⁵, and photogrammetry-based real-time capture systems) are expensive and can still be a bit fussy technologically speaking. These are two radically different forms, with different technologies, capacities, aesthetics and even ethical considerations. 360° video is the conceptual descendant of Robert Barker's 1787 panorama – a fixed visual asset in which the viewer can look around, and possibly trigger a few hot-spots.

Real-time capture systems, by contrast, enable interaction, with visual assets effectively being generated on the fly as the viewer moves around.

Augmented reality, technically speaking, is a subset of VR (a computer-generated emulation), but with an important twist. Whereas VR requires immersion into a computer-generated world, AR is an overlay on the world. That is, we see simulated and geo-located data or characters or images as well as the larger world. We can still easily interact with the world (and, depending on the



Augmented Reality through eye implants in *Sight* (Daniel Lazo/Eran May-Raz, ISR 2012)

system, with simulated artifacts), and AR lends itself to broader participation in the sense of allowing people to contribute their own virtual assets to the system (something much more difficult with VR). Technologically, AR goggles like Microsoft's HoloLens still have a long way to go, although mobile phones and tablets can also serve as 'portals' enabling users to see augmented overlays on the world. With VR, we leave the real world behind in order to enter a closed, simulated world; whereas with AR, we append a simulated layer onto the real world, and interact with both.

CINEMA: If VR allows us to «leave the real world behind», as you were saying, it has nonetheless a specific, though ambivalent relationship towards the documentary genres. In your text «The possible futures of documentary»⁶, you have reflected on this relationship, placing it into a historical perspective. Your text starts with the words: «History can be a great teacher, if only we put the right questions to it.» What kind of questions do we need to ask history in respect to Virtual Reality and documentary? What is new, what is familiar?

URICCHIO: Well ... Let's begin with «what's the reality claim implicit in VR?», «how has it developed?», «where is it headed?», and «with what ethical implications?» These are all questions that have been asked at one point or another of photography, film, and video, and they are questions that can help us understand the representational claims of these various media. Each medium has enjoyed moments when it was seen as nearly the equivalent of reality (usually to the medium's disadvantage – think of the decades during which film was dismissed as «the mechanical reproduction of reality» and the length of time it took to accept photography into art museums). These claims, in turn, laid the foundation for their later acceptance as the building blocks of various self-styled documentary movements, as mediated reality helped us «to see with new eyes» (Vertov). And more recently, these same claims were the stuff of critique as documentary kept pace with the world of post-structuralist, post-colonial, and even post-representational theory, and the apparatus with its structured power relationships was taken to task.

In terms of this rather compact historical arc, VR is still enjoying a naïve association with reality, although the fault lines of an impending critique are evident. But VR's reality claim has some twists. Whatever one thinks of the indexicality argument with regard to photochemical media (I'm a nay-sayer, but that's another argument), real-time VR-systems pose the issue in a new way. A system of measurements (the point cloud generated by a laser scan) has a pretty good claim to indexical status, but it's complicated by the algorithms that give it coherence, algorithms that are authored and require as much creative effort to emulate the rules of physics as to defy them. Or take the twist that is beginning

to emerge in the neuro-science community, where, some argue, real-time VR is processed in our brains as «experience» rather than as «representation», as is the case for film and television. If accurate, this too complicates the reality claim and how it is likely to unfold over time.

These issues matter for the documentary, at least as we've institutionally codified it in cinematic and televisual terms for the better part of a century. But they are complicated by other factors as well. Consider how we've tended – quite incorrectly as I've argued over the years – to dismiss as «naïve» the non-fiction films produced between 1895 and Grierson's narratively-loaded invocation of the term «documentary» in 1926. Is VR enjoying its «naïve» moment, as we slip into immersive states? And what are the implications of «being immersed» as opposed to «narrating» and «arguing»? Consider as well the ethical implications of representation in scenarios where the technologically empowered parachute in on the exotically disempowered. Yes, we have access to realities that we would otherwise miss, but at a cost and with implication. As I look at some of the VR documentaries that are being produced today, I get the sense that we have learned little from the neo-colonial vision of that first generation of ethnographic filmmakers.

CINEMA: Even though the modes of film production might sometimes follow relatively stable cultural paradigms, VR and AR seem to radically transform our practices and understanding of spectatorship. For classical cinema, the dominant model has long been the «spectator» – conceptualized as a passive viewer immobilized in the illusory realm of Plato's cave; VR seems to refigure this model towards concepts of a «user» who is able (and invited) to interact with the material. In what sense does the position of the viewer change in VR environments? And with whom or what exactly is s/he interacting when it comes to Virtual Reality?

URICCHIO: A user ... I like the word, especially for the agency that it seems to claim over the somewhat more passive sounding «receiver». I know that I'm on slippery ground here, thanks to the important work done by Hans Robert Jauss, Wolfgang Iser and several generations of reception researchers, but «user» sounds like an appropriate term for how I encounter books or films or whatever. One place where we can see the implications of how we imagine the human-text interaction is in the domain of narrative. Media studies is largely indebted to literary theory for its notions of narrative, and in that setting, narrative is a structured series of past events brought to life through the agency of a teller. «Pastness» in this case is a necessary condition of the written word, even if it unfolds in real time for the «receiver» who takes it in. Game studies have helped us to imagine a broader notion of narrative, one that is more experiential than textual. In games, we interact with an environment laden with narrative

elements and rules; multiple outcomes and multiple experiences are the norm, and indeed, there are generally no fixed resolutions. Like the fuller sense of the term «play» that we use to describe our interactions with games, narrative is the enactment of a coherent stance within an unfolding set of possibilities, it is precisely the «on the fly» experience we have when reading a novel, but without the fixity in the pages to come. Coherence is defined by the rule set, and «narrativity» is defined as the user's experience of negotiation and navigation within that rule set. This shift matters greatly for VR.

The key term for certain forms of VR – say, real-time capture as opposed to 360° video – is interaction, if by interaction we mean an encounter between the user and text that results in a reconfigured text. The distinction is an important one. 360° video is just that: video, a fixed asset. We can choose to attend to one part or another, but we can't change the video text any more than we can change the lines of a novel. Real-time capture systems, by contrast, enable unique views, enable the «user» to look behind things or wander around in space. It is more game-like in this sense. Each of these systems relies upon differing notions of narrative and differing notions of authorial agency.

«Authorial agency» is perhaps not the right term given its grounding in the traditional arts, but I mean here to point to the agents responsible for the existence of a text or textual environment before it comes into the hands of the user. In the case of 360° video, it is a familiar notion of authorship, not all that different from the video we've worked with for the past fifty years. But real-time capture is different. The maker constructs an environment and events within it that are bound by a rule system, within which the user is free to wander and interact. The environment is built of data points, and higher resolution images are created on the fly by algorithms responding to the user's visual field. Agency here is more diffused than in 360° video, including the designer of a world as well as the designers of the enabling algorithms. The onus is on the user to explore, to construct experience, and to render those experiences coherent.

This distinction is important because we can already see the next step of user-algorithm interactions coming in the form of eye-tracking headsets. It doesn't take much to imagine using pupil-tracked data as a navigational device, with the algorithms «anticipating» user interest and generating appropriate scenes, going far beyond their current deployment for «foveated rendering»⁷. In this scenario – and there are already working prototypes on the market – «the algorithm» (shorthand for a complex set of operations and models) makes choices extrapolated from user behavior, strengthening what we might consider the algorithm's programmatic «agency». The data sets that constitute these «environmental texts» permit different narrative paths, different experiences and points of view, and the question that arises is, who's the organizer of that path: the user, or the anticipatory algorithm and its designers? Although authoring algorithms re-

present a dramatic advancement of a principle, the principle itself is hardly a new one as we know from our Google searches and the recommendation systems deployed by Netflix, Spotify and Amazon. In each of these cases, algorithms make selections on our behalf (and that of the highest commercial bidder!), creating data sets that in turn enable our choices. They select and shape the data that we encounter, maintaining a pas de deux between «the user» and «the author», but diffusing that latter's agency among humans and responsive systems.

All this said, I would be remiss not to mention the danger – and indeed, the underlying purpose – of VR pupil-tracking systems. Yes, they can be used to do a lot of cool things, storytelling included; but they will be used to gather user data. They can track not only what one sees, but how the body responds based on dilations of the pupil. The same triggering system that paths a user through a narrative also generates mountains of data about social and consumption preferences. There is a steady creep in this kind of data acquisition, and we really need to call it out and draw a line about what we as a society think is appropriate and not.

CINEMA: You already mentioned that algorithms play a central role in current media environments. You have introduced the term «algorithmic turn», not only to pin down what is new about VR, but also for future examinations of film.⁸ Why do you think that algorithms constitute such a paradigmatic turn? In what ways do they change the relation between viewer, author, and the medium? And: Do we need a more radical rethinking of the terms in which we describe these relations?

URICCHIO: I think we are at a point where our existing vocabulary is inadequate to describe certain emerging phenomena. We are creatures of habit, quite sensibly informed by past experience and inherited categories of knowledge. Unfortunately, this can sometimes dull us to new conditions, which we temper by imagining through the old.

It may sound overblown, but I think that we are hovering on the edge of one of these shifts or ruptures. We are like those in the mid-15th century, about to experience a radical shift in the sense of self that would be expressed in the widespread acceptance of new representational technologies such as three-point perspective in rendering and the mass proliferation of the printed word. The algorithm, a system that has been with us since the ancients, has found new relevance in an era of big data (and all that era entails: digitization, connectivity, and formidable processing power). Algorithms are integral to this era, making their conditions of contingency (rather than certainty), their character as dynamic (rather than fixed), and their use for personalization (rather than standardization) relevant to an emerging epistemology. The world of finance, the supply and demand of information, goods, and energy, the markers of identity and citizenship, and so much more, are subject to algorithmic intermediation between

material conditions (transformed into data sets) and human agents. From these grand systems down to the smallest chips that enable our bank passes, digital cameras, and telephones to work, algorithms pervade our lives.

It's easy to make this sound sinister, but we need to recall that the algorithm is simply a tool. The human agents that design and deploy those tools merit our close scrutiny. But the tool itself can be used to enable new collectivities and broker the «wisdom of the crowd» that ensues. It can predict in useful ways; it can assess and complement; it can create ... In media terms, recommendation systems can assist navigation and ensure a «rich diet» of perspectives in a time of infinite choices, as easily as they can promote the interests of the highest bidder and trap us in an information bubble. They can construct texts and textual environments, like real time VR or the algorithmically generated stories that are finding a growing place in our newsfeeds. But they can as easily be used to filter, manipulate, and mislead. There's a problem of agency here, especially in a cultural order that rewards maximum accumulations of resource and power, since the rich and powerful will have undue influence in designing and deploying algorithms. But there's also an opportunity for making use of highly distributed knowledge and power.

Our traditions have not prepared us with a critical discourse for the dynamic, shape-shifting, and even personalized texts we encounter with, say, real time VR. We don't quite know how to account for the collaborative authorship of Wikipedia or the algorithmic agency of «Narrative Science»⁹. As we move from embodiment and fixity to a more ephemeral and contingent condition, we lack the critical language to constructively frame and assess these developments, other than by invoking the terms of the past. For now, we will have to make do with modifiers like «interactive» when talking about certain texts; or people like me will try to argue that existing terms like «narrative» have an expanded array of meanings in certain settings; or as you've suggested, we might turn to terms like «users» to describe what Jay Rosen called «the people formerly known as the audience». These examples are symptomatic. More fundamentally, we have to find ways to account for the role and workings of the algorithm as a third factor that has intruded in the classic binary of subject and object, and done so in a remarkably subtle manner. This is a new condition, the emergence of an epistemic era that has a different operating system than the philosophical order of the past 500 years has prepared us for. And that's why the 15th century is a reservoir of resonant experience.

CINEMA: *To what extent do those broader cultural shifts also shape our senses? Do you see a potential for VR and AR to eventually change the way we perceive the world?*

URICCHIO: I think that VR and AR will permit us to perceive new things about the world. The nature of that perception will to some extent be determined by how these technologies work: Do we process VR as «representation» or «experience»? If the latter, we can expect some perceptual shifts, and literature coming from the field of psychology suggests a spectrum of therapeutic applications claiming at least some correlations to change. It will to some extent depend on how VR-projects are packaged – as standalones, or as parts of larger media ensembles where messaging may come through other media conduits. And of course, it will to some extent depend on the results of our explorations of the expressive potentials of VR, for that is where the capacity for the truly new lurks.

AR, with its abilities for annotating and interacting with the world, poses a different set of possibilities, again for good or ill. The dangers of inappropriate or incorrect information are familiar, and in the case of AR, it won't just be about something, it will be on something. The dangers of distraction have taken a new turn thanks to the cell phone, and I'm not sure that AR will do more than contribute to the problem. And the promise of even more advertising – and right now, advertising constitutes the primary use case for AR! – is daunting. But on the positive side, AR offers the capacity to inform our travels in the world; to layer buildings and spaces with stories about their pasts and those of their inhabitants; to free documents and images from the archive and append them to the places to which they refer, offering the contextual advantages of space and time. These alone strike me as worth the risk.

To end on a final, dystopian note, AR and VR will both be informed by the same data-sucking project that has rendered digital media such a mixed blessing. As I suggested with VR pupil trackers, these systems are capable of sensing not only where we look, but thanks to the magic of pupillary dilation, how we respond – and this on a pre-conscious level. This level of informational granularity makes the data-traces that Google collects seem quaint, and requires a forceful rebuttal in terms of privacy norms. True, these norms change over time and across cultures, but for the moment, this is an area where it pays to tread slowly, hewing to established privacy norms until we are certain that a new regime of social contracts has been carved in stone.

-
- 1 André Bazin, «Le mythe du cinéma total», in: *Critique* 6 (1946), S. 555; repr. in: ders.: *Qu'est-ce que le cinéma?*, Paris 1958, S. 21–26.
 - 2 Bazin (wie Anm. 1).
 - 3 Cave automatic virtual environment (CAVE) is a virtual reality experience within a room-sized cube where projectors are directed to between three and six of the walls. It uses motion tracking to record and adapt to the user's real-time position (https://en.wikipedia.org/wiki/Cave_automatic_virtual_environment, zuletzt besucht am 11. 8. 2017).
 - 4 Lidar is a method to measure distances to a target by illuminating the target with a pulsed laser light (<https://en.wikipedia.org/wiki/Lidar>).
 - 5 Kinect is a motion sensing input device used for game consoles and PCs (<https://en.wikipedia.org/wiki/Kinect>).
 - 6 William Uricchio, «Things to come: the possible futures of documentary ... from a historical perspective», in: Judith Aston/Sandra Gaudenzi/Mandy Rose, *I-Docs. The Evolving Practices of Interactive Documentary*, London 2017.
 - 7 *Foveated rendering* is an upcoming video-game technique which uses an eye tracker integrated into a virtual reality headset to reduce the rendering workload by greatly reducing the image quality in the peripheral vision – outside of the zone gazed by the fovea (https://en.wikipedia.org/wiki/Virtual_reality_headset).
 - 8 William Uricchio, «The algorithmic turn: Photosynth, augmented reality and the changing implications of the image», in: *Visual Studies* 26/1 (2011), S. 25–35.
 - 9 Narrative Science is an artificial-intelligence technology company based in Chicago, Illinois, that invented Quill, an advanced natural language generation platform. Quill automatically creates a comprehensive story that attempts to meet the needs of the person it is intended for, using a combination of data analytics, natural language generation, and core artificial intelligence concepts (https://en.wikipedia.org/wiki/Artificial_intelligence).
-